Stochasticity and variability in the dynamics and genetics of populations

Perez de Vladar, Harold

Chapter 7

Perspectives: The evolution of quantitative characters under stabilizing selection, mutation, and drift.

7.1 INTRODUCTION

In getting to understand the nature of selection, some contradictions arise, at least from the classical theory of population genetics with respect to maintenance of genetic variability. Not only that we know, empirically and theoretically, that selection depletes genetic variance, and that mutation, with the aid of recombination induce it, but also that the levels that we expect from theory are in most cases lower than those experimentally observed (Maynard-Smith, 1983; Charlesworth et al., 1982, and references therein). But the question is not only one of the matter of scale, but also of logical thought. A standard argument is that extreme phenotypes tend to be less viable than intermediate ones. Examples (a) giraffes with too short necks are dissadvantaged in that they cannot reach the food composing their main diet, and the blood pressure might be too high in their head's vascular system when they bend to drink. Conversely, giraffes with too long necks may not have enough blood pressure in their heads and brains at normal postures; thus intermediate sizes are of higher viability. (b) The amount of chlorophyll in higher plants when too low, cannot account for enough energy transduced for the individual's vital processes; too much chlorophil requires an even bigger amount of energy investment for creating cellular, anatomical and physiological structures that are neither compensated by the energetic gain, nor useful because it would imply an increased metabolic rate that might not be sustainable by size and CO_2 intake rates, and inability to dissipate heat. We can build this kind of argument untill the (z) with most phenotypic characters of any species. For instance, the stabilizing selection hypothesis seems a widespread possibility in natural populations.

Where is then, the contradiction? Although we would expect this kind of selection to be observed frequently, studies on patterns of selection in the wild have revealed that it is scarcely present, and that directional selection is the conspicuous choice (Kingsolver et al., 2001; Hoekstra et al., 2001). If this is the case, there is excess of genetic variability attributable to mutation and linkage. But mutation and recombination rates are not high enough as to account for such levels of variability.

A sound hypothesis is that most characters are under pleiotropic effects, and that these characters have opposing effects to the traits under selection. This alternative, apparent stabilizing selection, was explored in chapter 6.

Another reasonable alternative is that indeed weak stabilizing selection is acting. If the mean phenotype is displaced from the optimal state, directional selection effectivelly acts towards the new optimum. In this case genetic variability can be more easily argued to be maintained at higher levels, as explained above. Notice that the response of genetic variance to selection over a trait tends to be delayed in prolonged application of selection, in particular if there are pleiotropic effects. Unless deviations would be very far from the optimum in one specific (genetic) direction, for example if drift would introduce an irrationally large deviation, genetic variance is unlikely to be (statistically speaking) changed. The return to the optimum would proceed with a linear rate, with variance essentially unchanging at noticeable scales, and experiencing directional selection.

This second alternative, is compatible with the experimental observations (Hoekstra et al., 2001, although with secondary importance after directional selection following meta-analyses)) and with the logical argumentations about the lowered fitness of the extreme phenotypes.

The evidence and arguments for the stabilizing nature of selection, thus demands a dedicate analysis. The approach of the

statistical mechanics theory developed in the previous chapters, will be extended to this situation. First to be able to quantify the course of evolution, and second to make available the quantitative tools for a comprehensive comparative evaluation of the possible evolutionary forces in action.

But the stabilizing selection situation, in inherently complicated. Any given equilibrium between stabilizing selection, mutation, and drift (SSMD) would have many possible microscopic equilibria (Barton and Shpak, 2000; Turelli and Barton, 2004). The dynamic of two or more loci coupled through a trait under stabilizing selection leads to a range of possible dynamics, that are far from entirely characterized (Barton and Shpak, 2000; Willensdorfer and Bürger, 2003; Gavrilets and Hastings, 1993). Although at the moment we are not directly concerned with these dynamics, and the characterization of the local equilibria, it is clear that perturbations to the allele frequencies (e.g. by drift) can induce metastability in these states Barton and Rouhani (1987); Rouhani and Barton (1987). Thus continously perturbating the equilibrium states, leads to an ever changing microscopic dynamics that show increased genetic variance from quantitative measurements approach.

The situation is not as trivial as in directional selection. There are non-linearities in the dynamics, because selection occurs over the squared mean trait, which among other consequences, it fully couples the loci. Hence a decomposition of a polygenic trait as a many independent one-locus problems is not possible, as it was in the case of directional selection. Yet there are stratagems to be victorious in averaging out the microscopic variables. Maximum entropy could work if we find appropriate macroscopics. In short, the question is whether the local equilibrium approximation holds. If it does, we would be free to track evolutionary dynamics, like moving optima, enhaced (or relaxed) strength of selection to the extremes, etc.

7.2 MAX-ENTROPIC APPROACH

In the chapter 3, the methodology inspired by the analogy with statistical mechanics in physics was derived and applied in detailed analyses to directional selection, to study the evolution of quantitative characters in univariate traits. In chapters 4-6 this methodology was extended for multivariate polygenic traits with unequal effects.

We employed maximal entropy as a starting point, constrained by the macroscopic variables that are maintained by the evolutionary processes. That is the trait -or fitness- (maintained by selection), and genetic variability (maintained by mutation), if the selective scenario is directional over an additive trait.

Stabilizing selection over a quantitative character removes from the population those genotypes whose traits are far from an optimum. This is equivalent to have directional selection against the genetic variance. But this is not enough, since we would get for equilibrium a single point at trait zero, without any variance. So we must also include genetic variability (to account for mutations, if the rate is $\mu > 1/4N$, and directional selection over the trait, towards an optimum. Intuitively, this should be enough. But since selection is assumed over the mean trait, then fitness of the mean trait is not the same as mean fitness, so a second order term would appear (variance of the mean trait). Mathematically, we typically choose to model stabilizing selection as a gaussian landscape of fitness:

$$W_z := \exp\left[-\beta(z - z_{op})^2\right] \tag{7.1}$$

We first average over the frequencies of the traits ($P(z)$ to get the mean fitness, $\bar{W} = \int W_z dP(z)$. Now if β is sufficiently small (selection over the trait is weak) we can expand to get

$$\bar{W} \simeq \exp\left[-\beta\nu_z - \beta(\bar{z} - z_{op})^2\right] \ . \tag{7.2}$$

Notice that when expanding the square in the parenthesis, three terms appear: a constant (βz_{op}^2), $\beta z_{op}\bar{z}$ that is directional selection towards the optimum, and $-\beta\bar{z}^2$ that is selection against the squared trait. The stationary distribution of the allele frequencies is recovered from entropy maximization (Eq. 3.6, Ch. 3) constraining the expectations of the above quantities, i.e.:

- Normalization of the distribution.

$$\int_{(0,1)^n} \psi(\mathbf{p}) d^n\mathbf{p} = 1 \rightsquigarrow \lambda \rightsquigarrow \mathbb{Z} \tag{7.3}$$

- Selection of the mean trait towards the optimum

$$\int_{(0,1)^n} \bar{z}\psi(\mathbf{p}) d^n\mathbf{p} = \langle \bar{z} \rangle \rightsquigarrow 2N\beta \tag{7.4}$$

- Selection against genetic variance.

$$\int_{(0,1)^n} \nu_z\psi(\mathbf{p}) d^n\mathbf{p} = \langle \nu_z \rangle \rightsquigarrow 2N\sigma \tag{7.5}$$

- Selection against the variance of the mean trait

$$\int_{(0,1)^n} \bar{z}^2\psi(\mathbf{p}) d^n\mathbf{p} = \langle \bar{z}^2 \rangle \rightsquigarrow 2N\alpha \tag{7.6}$$

We know that maximizing entropy leads to the distribution 3.6, that in this case of SSMD case it would be

$$\psi(\mathbf{p}) = \frac{\phi}{\mathbb{Z}} \exp\left[2N\beta\bar{z} + 2N\alpha\bar{z}^2 + 2N\sigma\nu_z + 2N\mu U\right] \tag{7.7}$$

with $\phi \equiv \phi(\mathbf{p}) = \prod_{\ell=1}^{n}(p_\ell q_\ell)^{-1}$, as explained before, and if we are able to compute the integral $\mathbb{Z}$, we will have a macroscopic description of the system, that is supported by and consistent

with the microscopic stationary dynamics. But more interesting than finding the closed form expression (if possible at all) is to be able to predict, also from this macroscopic point of view, the dynamics. Microscopically, this is given either in a stochastic version, or probabilistically by the corresponding Wright-Fisher process, and its diffusion equation, respectivelly. Actually those will be our points of comparison.

The partition function can be explicitly written if we define how the trait relates to the genetic variables. As before, we assume an additive trait of n loci, each with constant effect γ_ℓ (constant in the sense that they do not evolve, but each effect is in general different at every locus). As treated in previous chapters (see also appendixes) the mean trait and genetic variance are functions of the allele frequencies. However, the squared trait introduces some tricky properties that complicate the computation of $\mathbb{Z}$, since it cannot longer be expressed as the product of the partition functions of independent loci. So, for reasons that will be obvious later, I will express any power of $\bar{z}$ implicitly. Thus the partition function is

$$\mathbb{Z} = \int_{(0,1)^n} d^n\mathbf{p} \prod_{\ell=1}^{n} (p_\ell q_\ell)^{-1} \times \tag{7.8}$$
$$\times \exp\left[2N\beta\bar{z} + 2N\alpha\bar{z}^2 + 2N\sigma \sum_{\ell=1}^{n} \gamma_\ell^2 p_\ell q_\ell + 4N\mu \sum_{\ell=1}^{n} \log(p_\ell q_\ell)\right] ,$$

which cannot be computed in a closed analytical form. Before dooming the expression 7.8 to numerical computations, we can do something about it. The reason is that as it is, its computation is $n-$dimensional, thus prone to slow convergence, and since the allele frequencies might be clustered, the integrals might be close to zero. Also, there are some simplifications that can be made. The whole point, is that if we are able to compute $\mathbb{Z}$, we can as well compute any statistic. Hence we are

able to test the local equilibrium as a model for the polygenic evolutionary dynamics.

The trick, following Barton (1989, p. 64), is to transform the $n-$dimensional integral into a complex-valued 1-dimensional integral, which if not anallytically solvable, at least simplifies (a) the calculation of the expressions of the traits, and (b) their numerical computations.

I will give only a sketch on how to proceed with the calculations. The intermediary steps should be straightforward. Lets calculate a function $F(\bar{z})$. It can be expressed as an integral of a Dirac delta function as:

$$\begin{aligned} F(\bar{z}) &= \int_{-\infty}^{\infty} F(\zeta)\delta(\zeta - \bar{z})d\zeta \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(\zeta)\exp\left[-i\omega(\zeta + \bar{z})\right]d\zeta d\omega \ . \end{aligned} \tag{7.9}$$

where the second expression employed the (inverse) Fourier transform of an exponential function to express Dirac's delta. If the last expression is not disregarded by the reader, the advantages are clear: the function F is evaluated in a mute variable, ζ and the dependence of $\bar{z}$ are segregated to the exponential factor, which avoids cross terms (like $p_i p_j$) of the allele frequencies. Following this formula 7.8,

$$\begin{aligned} &\exp\left[2N\beta\bar{z} + 2N\alpha\bar{z}^2\right] = \\ &\sqrt{\frac{\pi}{|2N\alpha|}}\int_{-\infty}^{\infty}\exp\left[-i\omega\bar{z}\right]\exp\left[-\frac{(2N\beta - i\omega)^2}{8N\alpha}\right]d\omega \end{aligned}$$

This expression can be substituted into Eq. 7.8, which after some calculus it gives

$$\mathbb{Z} = \sqrt{\frac{\pi}{|2N\alpha|}}\int_{-\infty}^{\infty}\exp\left[-\frac{(2N\beta - i\omega)^2}{8N\alpha}\right]\prod_{\ell=1}^{n}\mathbb{Z}_{\ell\ (\mu;-\frac{i\omega}{2N}\gamma_\ell;\sigma\gamma_\ell^2)}d\omega \tag{7.10}$$

where the per-locus partition functions are given by:

$$\mathbb{Z}_\ell := \int_0^1 \exp[2N\beta(2p_\ell - 1) + 4N\mu \log(p_\ell q_\ell) + 2N\sigma p_\ell q_\ell](p_\ell q_\ell)^{-1} dp_\ell \tag{7.11}$$

Admittedly, the expressions are not too simple, and lack general solutions[1]. Even if each $\mathbb{Z}_\ell$ is to be integrated numerically, these are independent among each other. Thus originally, there were n fully coupled integrals (big problem), and now there are $n+1$ integrals, one of which depends on the other n integrals, but which are independent among each other (small problem).

The general partition function. Eq. 7.10 can be dissected into two terms, generally of the form

$$\int F_{1\ (\omega|\alpha,\beta)} F_{2\ (\omega|\mu,\sigma)} d\omega$$

that is, it is separable into two terms which depend on non-overlapping (intensive) variables. This is a very useful property when calculating the macroscopics.

Take notice that solving a one-locus problem for stabilizing selection might sound as a trivial exercise. But there are three reasons why at the moments it is desirable to do it. First, the polygenic expectations can be represented (exactly) as convolutions of the one-loci corresponding formulas (see appendix D.2); thus it is a necessary move. Second, the Fourier-representation, being equivalent to the n-loci representation of the partition function (and of the expectations), can be checked for one locus problems; this is just a control of the numerical experiments.

[1]The per- locus partition function, however can be expressed in Taylor series over $\sigma = 0$, in which case the formula for directional selection is recovered. The terms for $\sigma^k,\ k > 1$ involve derivatives of this partition function at $\sigma = 0$, and thus are the single-locus-directional-selection expectations.

Third, the one-locus SS problem has implications for the DS case, with respect to the boundary problem that appear near $N\mu = 1/4$, and makes the local equilibrium fail. One locus SS might provide a subterfuge for this collapse.

7.3 SINGLE LOCUS DYNAMICS

In order to solve the general case of the multi-locus dynamics, we need to have a complete characterization of the statistics of single locus model. Among virtues of a mean trait affected by only one locus is that the variance of such mean trait is proportional to the genetic varince, since

$$\begin{aligned} \bar{z}^2 &= \gamma^2[2p-1]^2 \\ &= \gamma^2 - 2\nu_z \ , \end{aligned}$$

Because the effects γ are non-evolving parameters, the statistical mechanics in this case does not require constraints in both $\bar{z}^2$ *and* ν_z; it rather requires constraints in *one* of them. Yet the general locus formula 7.8 applies. For one locus, we will end up with only one of these two quantities, say genetic variance, and the multipliers to the constraints over entropy maximization will be reduced as

$$\begin{aligned} \sigma &\mapsto \sigma - 2\alpha \\ \lambda &\mapsto \lambda - \alpha\gamma^2 \\ \mathbb{Z} &\mapsto e^{-\alpha\gamma^2}\mathbb{Z} \ . \end{aligned} \tag{7.12}$$

Hence, a single-locus model of SS requires only three macroscopics. We could say that this case is a small extension to the directional selection case, where we took a second order approximation to mean fitness (which would result in including genetic variance as a second order correction term in the mean

fitness term). But his extended model gives whole new properties, and is more than a small quantitative correction. First, on the technical side, including selection against the variance does not allow the privilege of having closed form solutions of the partition function (or of the expectances), so we must proceed numerically and/or with some approximations. Notwithstanding, the integration procedures are not too demanding computationally, since most integrands are well-behaved, even near the point $\mu \sim 1/4N$. Second, including selection against (or for) genetic variance breaks the symmetry that exists in directional selection with respect to $\langle \bar{z} \rangle$, where this function is odd with respect to β and $\langle \nu_z \rangle$ is even. Directionally selecting for a favorable allele is -from the point of view of genetic variance- equivalent to selecting for the contrary allele. Also the measure $\langle U \rangle$ would be unaffected. However, if selection over genetic variance is included all these symmetries disappear. Selection will still deplete genetic variation, but on one direction (favoring an allele) will in general be higher than when favoring the contrary allele. The same is true for generic variability $\langle U \rangle$.

The evolutionary dynamics of this one-locus system under SSMD can be computed through the local equilibrium approximation (Section 3.2). That is to calculate the rate of the effective parameters $(\mu^*, \beta^*, \sigma^*)$ that correspond to the quantitative measurements ($\langle U \rangle$, $\langle \bar{z} \rangle$, $\langle \nu_z \rangle$) at every time-point during transient (i.e. non-equilibrium) evolution. Mechanistically, this is determined by the change of the allele frequencies in the population, averaged over the drift realizations. This is described by the diffusion equation (Crow and Kimura, 1970, Ch. 8). The change in the observables (the quantitative variables required for the max-entropic restrictions) follows Eq. 3.13. Thus in addition to the mutational variability U and the mean trait $\langle z \rangle$, we also need to include as observable the expectations of the genetic variance. To proceed in local equilibrium analysis (Eqns. 3.15

and 3.16), the matrices of genetic co-variances B and of fluctuations C are required.The matrices are, respectively:

$$B = \left\langle \begin{array}{ccc} H & -2\bar{z} & 2\bar{z}^2 \\ -2\bar{z} & \nu_z & -\bar{z}\nu \\ 2\bar{z}^2 & -\bar{z}\nu & \bar{z}^2\nu \end{array} \right\rangle \tag{7.13}$$

$$C = \{\text{Cov}(A_i A_j)\}_{i,j\in\{U,\bar{z},\nu_z\}} \ , \tag{7.14}$$

where each of the terms can be found in the appendix D.2, and the dynamics are given by

$$\frac{d\alpha^*}{dt} = C^{-1}.B.(\alpha - \alpha^*) \ . \tag{7.15}$$

Here $\alpha = (\mu, \beta, \sigma)^\top$ ($\top$ stands for transpose, since the formulas require a column vector). Beware that even if this equation seems to 'include' Eq. 3.24 (for directional selection) in that there is an extra column with respect to that, the explicit forms of the statistics are different. Implicitly there is a resemblance, although the specific forms of the expectances would be very much different, because -as can be noticed from Eq. 7.11 and the formulas in the Appendix D.2-, the same statistics will have different quantitative values in directional and stabilizing selection.

Now it is possible to proceed for some case studies. Compare to directional selection, where only two intensive variables existed: mutation rate and selective value of the trait. Now there is another one, selective value of the genetic variance. Thus it is possible to have a bigger scope of possibilities in the evolutionary dynamics. Only a handful will be investigated here.

Evolution towards directional selection, mutation and drift. This single locus model, would reduce to that of directional selection if we let $\sigma = 0$. How would be the response to this situation?

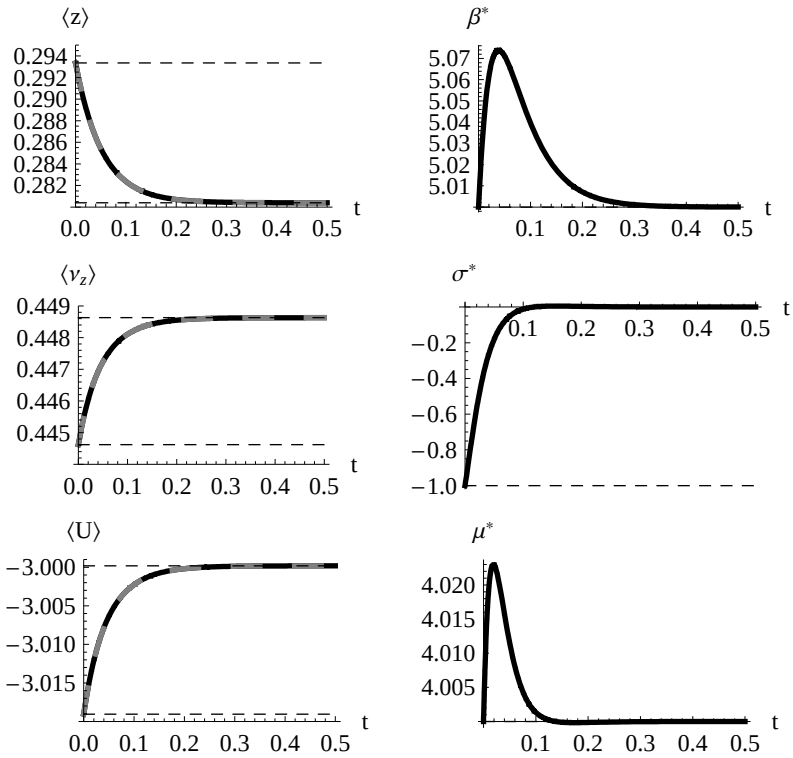

Figure 7.1: Evolutionary dynamics from an initial equilibrium maintained by selection for the trait, selection against the variance, mutation and drift, evolving towards a state without selection against the variance. This is parametrized by $(N\mu, N\beta, N\sigma) = (4, 5, -1) \to (4, 5, 0)$. Black curves: local equilibrium calculations; gray dashed curves: diffusion equation integrations; dashed thin lines: equilibrium values.

Fig. 7.1 shows this experiment. As we would expect, the genetic variance increases (since there is no selection against it anymore). Other consequence is increased mutational variability and reduced mean traits. But notice that even when the

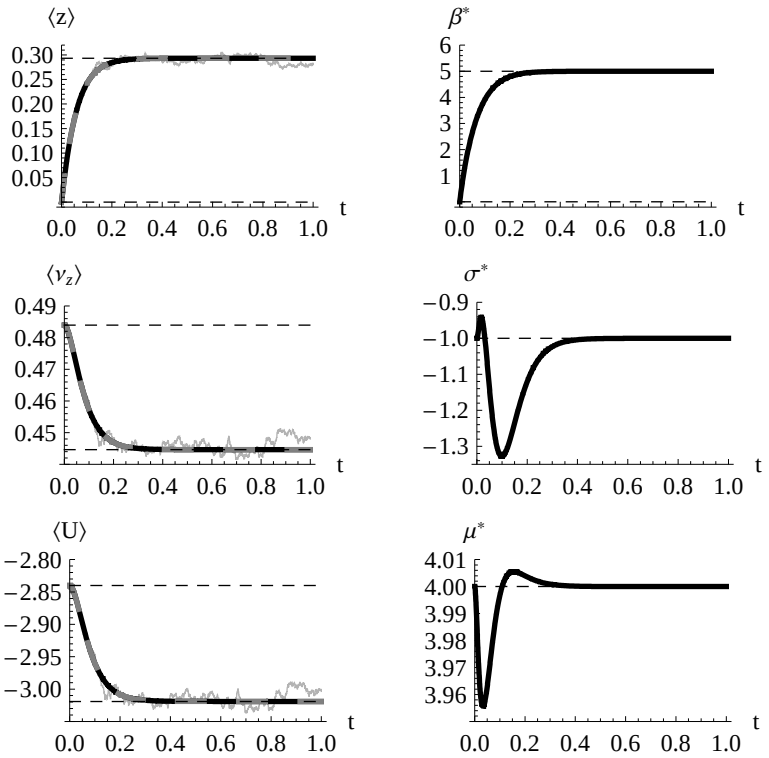

Figure 7.2: Evolutionary dynamics responding to a shift in the optimum. Starting at an equilibrium state where the optimum value is at the origin, the system evolves towards an equilibrium at a higher optimal value. The system is parametrized by $(N\mu, N\beta, N\sigma) = (4, 0, -1) \to (4, 5 - 1)$. The thin gray lines, are averages over 600 realizations of solutions to the Wright-Fisher model, with the above parameters with a population size of $N = 10$. Otherwise as in Fig. 7.1.

traits are (in average) reduced, they are more widely spread, so we would expect to find bigger and lower extremes.

Shifted Optimum. Many problems in stabilizing selection assume that there is a shifting optimum that motors the evolutionary dynamics. This situation can be modeled by a sudden change in this optimum, to which the response will smoothly described by the local equilibrium. The trait will experience directional selection. As regarded in Fig. 7.2, the response is comparable to that of directional selection (Figs. 6.2 -6.3): increase in the trait, and depletion of genetic variance and mutational variability.

Although at first sight no major qualitative difference is noticed with respect to the DSMD, recall the problem near the border $\mu \sim 1/4N$ (Fig. 3.11, and section 3.4) where allele fixation suddenly becomes abruptly likely and makes local equilibrium inapplicable. If genetic variance is 'controlled' (i.e. prone to selection), the problem seems to disappear. Figure 7.3 shows that shifting the optimum at different mutation rate, closer to $1/4N$ works out well. The statistical mechanical approach does not fail like in DSMD. (Actually, although only a matter of numerical methods, the differential equations of the local equilibrium approach behaves better than the partial differential equations of the diffusion equation which has, under many methods, a leak of probability density mass. This was corroborated through averaging 600 realizations of the corresponding Wright-Fisher process; Figs. 7.2 and 7.3.)

Deaccelerating the mutation rate. The '$4N\mu$-boundary-problem' in the DSMS formulation with statistical mechanics limits the case studies that we can address with this method. That is unfortunate because the effects of lowering (or increasing) the rates of mutations or studying bottlenecks, for example, cannot be analyzed. However, this problem seems to be absent under SS. Fig. 7.4 shows how evolution, using the top-down approach, is accurately described. Mutation $N\mu$ was switched

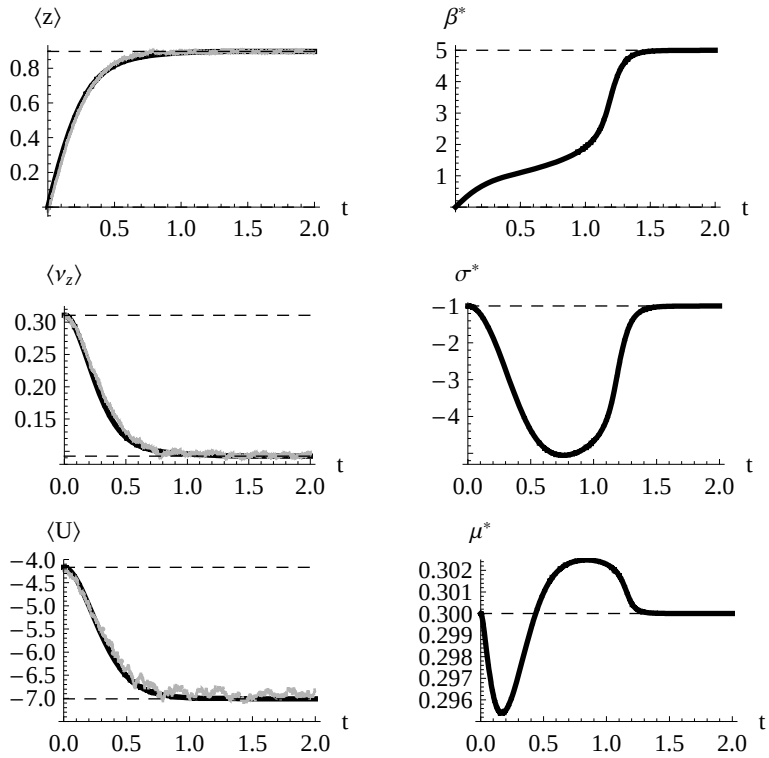

Figure 7.3: Evolutionary dynamics responding to a shift in the optimum at low mutation rates, with $N\mu = 0.3$. Otherwise as in Fig. 7.2

down to 0.3 without apparent discordance in any of the macroscopics. Notice that the local variables show curious paths as $N\mu \to 1/4$.

Revisiting directional selection. The paragraphs above show that as $N\mu \to 1/4$, where statistical mechanics fails under DSMD, evolution is authentically traced. Moreover, when

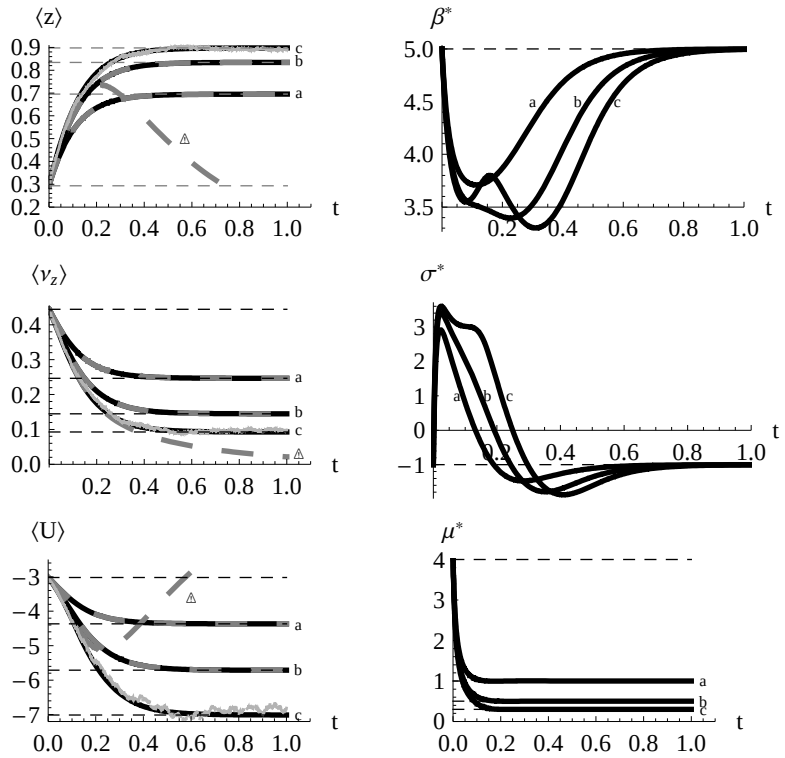

Figure 7.4: Evolutionary dynamics when mutation is slowed down (a) $N\mu = 1$, (b) $N\mu = 0.5$ and (c) $N\mu = 0.3$. Legends as in Fig. 7.2. Regard that for (c) integrations using the diffusion equation have a leak of probability density (⚠).

$\sigma \to 0$ the statistics of DSMD are recovered. Thus there is a new prospect: if at equilibrium we constrain $\sigma = 0$, but we allow it to evolve, we could predict the dynamics of DS near the critical mutation rates. I performed this experiment considering the SSMD statistics at $N\mu = 0.3$. Naturally at equilibrium DS and SS produce the same results. But in the course of evolution,

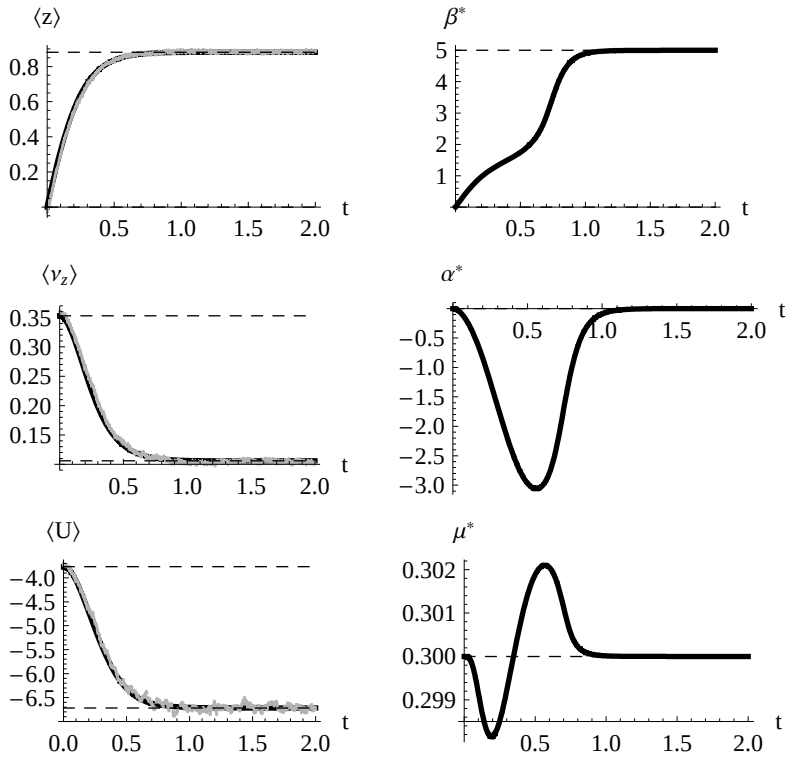

Figure 7.5: Evolutionary dynamics after shifting the optimum at low mutation rates. The solutions are compared to the Wright-Fisher process. Parameters of the system: $(N\mu, N\beta, N\sigma) = (0.3, 0, -1) \rightarrow (0.3, 5, -1)$. Otherwise, as in Fig. 7.2

as shown in Fig. 7.5, the predictions match that of the Wright-Fisher model. Surprisingly, the method is robust: calculation at the critical point (setting $N\mu = 1/4$) still give propitious predictions, Fig. 7.6 (curiously, the computing time for the numerical solution is considerably higher than in other cases, roughly an hour, at least two orders of magnitude higher than for the

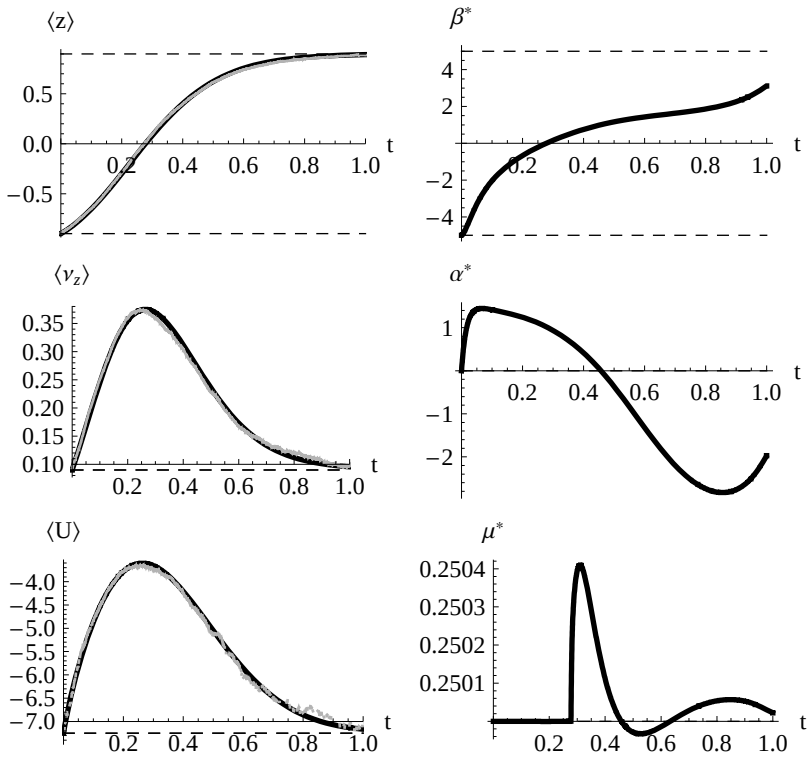

Figure 7.6: Evolutionary dynamics after shifting the optimum at the critical mutation rate. The solutions are compared to the Wright-Fisher process. Parameters of the system: $(N\mu, N\beta, N\sigma) = (1/4, 0, 0) \rightarrow (1/4, 5, 0)$. Otherwise, as in Fig. 7.2

previous ones, between one and two min.)

Summary. The one-locus model has shown that (i) the SM method is not limited to simple directional selection descriptions, (ii) the point at $4N\mu = 1$ is not a limitation for the macroscopic dynamics when including genetic variance as constrained macroscopic, (iii) there is a smooth limit from stabilizing selection to directional selection when $\sigma \to 0$ and (iv) that evolution under directional selection might have stages where the effective forces out of equilibrium are of stabilizing nature.

7.4 FORMULATING POLYGENIC DYNAMICS

Much of the work for polygenic systems under stabilizing selection has been achieved using a 2-locus model or a haploid approximation (Bürger, 2000, Ch.VI). Weak (Gaussian) stabilizing selection is unable to maintain variability, provided that the contribution of all loci over the trait is the same (Wright, 1935, unless linkage is strong Karlin and Feldman, 1970). However, if the allelic effects at each locus are different, then it is possible to keep elevated genetic variance Nagylaki (1989); Gavrilets and Hastings (1993). In any case, there are many microscopic equilibria. For a system as simple at two loci, there are at least 18 equilibrium points (Willensdorfer and Bürger, 2003). All of these are able to maintain distinct degrees of variability. Yet they do not account for the levels observed in quantitative traits, for typical allelic mutation rates.

The other extreme, which is also a common approach, is the infinite alleles model. As we mentioned before, the House of Cards (Kingman, 1978; Turelli, 1984) and the Gaussian approximations (Kimura, 1965a; Lande, 1976) can give an idea on the amount of quantitative variation that is maintained by mutation-selection drift. But as Slatkin and Frank (1990) point, "neither model can be regarded as being typical". The amount

of variation that is predicted curiously depends very slightly on the amount of loci (Turelli, 1984). This is a consequence that the distinct loci have alleles are close to fixation, and variability is maintained in only one of them.

The exact model (hypergeometric), on the other hand, consists of many genes of equal effects, and allows to identify many possible combinations of microscopic equilibria (Barton and Shpak, 2000). Essentially, all genotypes with the same amount of favorable alleles have the same fitness (although not all of them are stable).

We thus see, that to analyze the composition of the alleles in the population requires a thorough characterization. Yet the macroscopic states to which these states correspond are much more simpler. As mentioned above, and seen from the partition function 7.10, the polygenic statistics are not simple 'superposition' of the effects of each locus. Yet the polygenic framework for SSMD relies on the properties of single loci, although in in a non-linear way. Since we need four macroscopics to define the SSMD equilibrium, then the matrices B and C require also an extra dimension, the statistics for $\bar{z}^2$. The problem is more than just calculating the necessary parameters in these matrices. Not only that we need to calculate for each locus these quantities, but we need to convolve them with the trait distribution in a complex space. To my big regret there is little hope that analytic expressions are possible; although perhaps approximations will be workable, which certainly would enlighten our understanding of evolutionary quantitative genetics. For the moments the goal is to set up the problem. Extensive investigation of the macroscopic solutions is needed, since the microscopic dynamics have a wide range of solutions whose consequence over the macroscopics' we don't know.

For this purpose we can apply the theory stated above, which accounts to extend the matrices 7.13 to include the effects of

the variable $\langle \bar{z}^2 \rangle$. This leads to

$$B = \left\langle \begin{array}{cccc} H & -2\bar{z} & 4(\nu_{\max} - \nu_z) & -4\bar{z}^2 \\ -2\bar{z} & \nu_z & \mu_{z\,3} & 2\nu_z\bar{z} \\ 4(\nu_{\max} - \nu_z) & \mu_{z\,3} & \mu_{4\,z} & 2\bar{z}\mu_{z\,3} \\ -4\bar{z}^2 & 2\nu_z\bar{z} & 2\bar{z}\mu_{z\,3} & 4\nu_z\bar{z}^2 \end{array} \right\rangle \tag{7.16}$$

$$C = \{\mathrm{Cov}(A_i A_j)\}_{i,j\in\{U,\bar{z},\nu_z,\bar{z}^2\}} \tag{7.17}$$

where each of the terms can be found in the appendix D.2.

The reader may notice the following difference with respect to 7.13. Besides the above mentioned extension, the column and row corresponding to ν_z have been written in different way. The reason is that some identities do not apply for multilocus formulas. For example, the term $2\gamma^2 pq\gamma(1-2p)$ for one locus corresponds to $-\nu_z\bar{z}$ however the term $\sum_\ell 2\gamma_\ell^2 p_\ell q_\ell \gamma_\ell (1-2p_\ell)$ is *not* the same as $-\nu_z\bar{z}$, but rather the third moment, $\mu_{3\,z}$ of the trait within a population. For one locus is then true that $\mu_{3\,z} = -\nu_z\bar{z}$ in the same way that it is true that, as we saw, $\bar{z}^2 = 1-2\nu_z$. But these identities are not extendable to the polygenic formulas. In other words, the statistics for the mean trait are not necessarily a lumping of the statistics of the individual loci, as in the case of the mean trait or of DS.

We need to calculate these macroscopics numerically. There are some tricks to calculate them, from the Fourier-space integrals, as given in Appendix D.2. In short, separating the partition function as indicated above leads to some ways of expressing the polygenic statistics as a function of the single-locus statistics. These calculations are much simpler than those in the space of genetic frequencies, essentially because these are 1-dimensional calculations. Still, the amount of time they take to compute is enormous , making it impractical to compute for many loci (see below). Thus further work is needed to advance

in this direction. But for the moments, we are able to make some equilibrium predictions.

If we set $d\langle A_j\rangle/dt = 0$ then we obtain the conditions for mutation-selection-drift equilibrium. For the sake of simplicity, lets assume that the trait distribution is normal. In that case $\mu_{z\ 3} = 0$ and $\mu_{z\ 4} \propto \nu^2$ (as in Barton and Turelli, 1987, following Lande, 1976). We can then obtain some expressions which we are able to interpret. For the genetic variance we have that

$$\langle \nu_z \rangle = \frac{2\mu}{\beta}\langle z \rangle - \frac{2\alpha}{\beta}\langle \nu_z \bar{z} \rangle \ . \tag{7.18}$$

The last term is absent in the case of directional selection. Notice that even if $\mu \to 0$ the variance is still maintained by selection. Thus it is possible to increase the genetic variance without increasing the rate of mutation. This effect was identified by Gavrilets and Hastings (1993) for a two locus model. But we find here for arbitrary number of loci. The relation between the expectancies presented here and previous estimates of genetic variance, like those mentioned above, is that the expectancies discussed here is comprised of all those microscopic equilibria. Since drift is present, there will be shifts between those microscopic equilibria, and we are averaging over those.

Figure 7.7 shows that *without changing the mutation rate*, genetic variance can be maintained, and even increased by other factors. The observables are intrinsically pleiotropic, thus the change of a given factor results in the change of all observables. This is what is show in Fig. 7.7: when we select for the trait (A,B), for the genetic variance (C,D) and for the variance of the trait (E,F) we still get a response of the genetic variance, in the last two cases, an increase. These changes however, are lamer when more loci are present (even if most of them are of small effect), also indicating a stronger pleiotropic effect.

Although at the moments the efficiency of the algorithms are

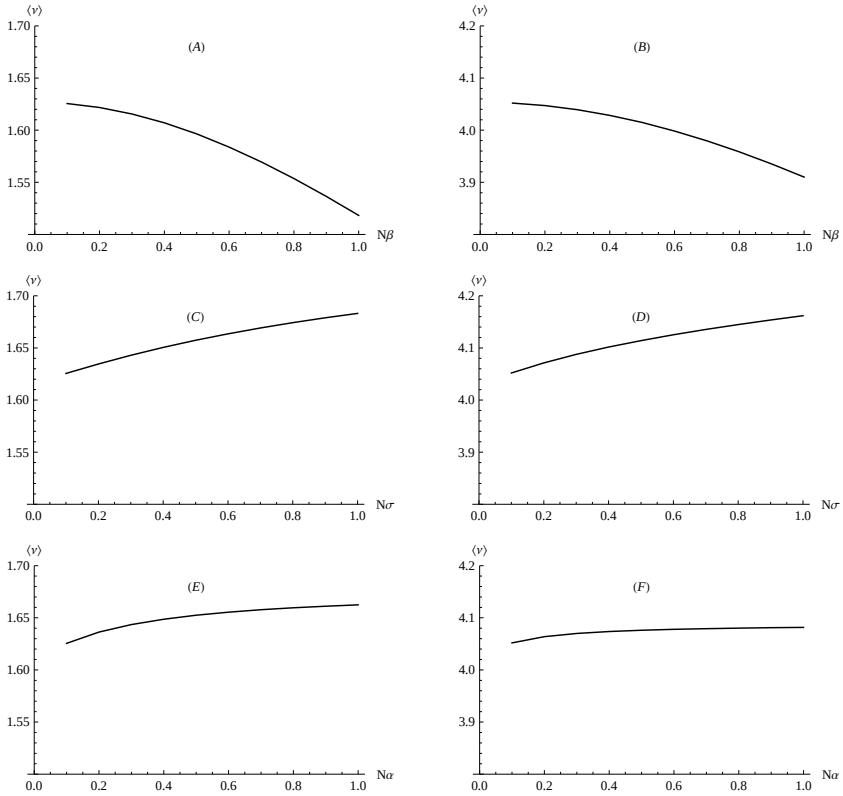

Figure 7.7: Expectancies of the genetic variance of polygenic traits under stabilizing selection as a function of (A,B) $N\beta$, (C,D) $N\sigma$, (E,F) $N\alpha$. Left column: traits composed of 4 loci of effects $(0.15, 0.76, 1.09, 1.73)$. Right column: traits composed of 10 loci of effects: $(0.03, 0.12, 0.36, 0.50, 0.68, 0.80, 0.83, 1.01, 1.38, 2.6)$. Unless the parameters are changed as indicated in the axes, these are $N\beta = N\sigma = -N\alpha = 0.1$ and $N\mu = 0.5$.

limiting the numerical analyses, I calculated the evolutionary response for a two locus system after shifting the optimum 7.8. Here we recover the classical result that selection depletes ge-

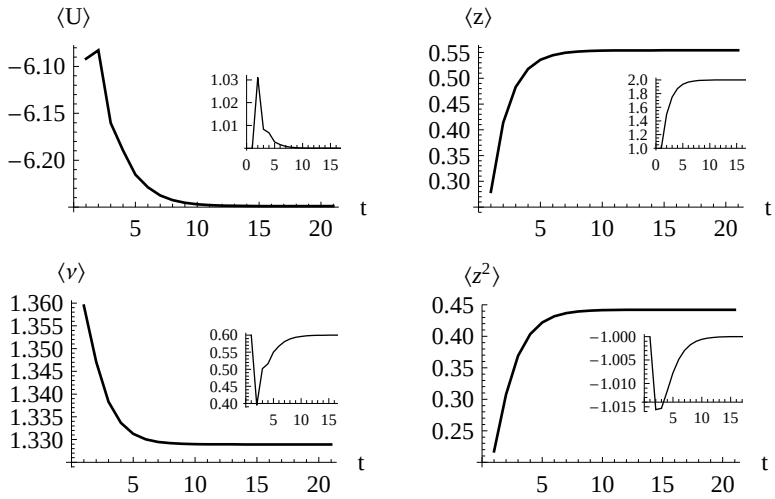

Figure 7.8: Evolution of a tri-locus trait under stabilizing selection. Evolutionary response of (A) genetic variability (inset: local mutation rate $N\mu^*$), (B) mean trait (inset: local selection $N\beta$), (C) genetic variance (inset: local selection $N\sigma$), (D) square of the mean trait (inset:local selection $N\alpha$). The initial state is given by $N\mu = 1.0, N\beta = -1.0, N\sigma = 0.6, N\alpha = -1$, (equivalent to the optimum at $z_o = -1/2$), and $N\beta$ is changed to 2.0 (equivalent to a shift in the optimum to $z_o = -1$). The loci have equal unit effects.

netic variance. However, we re not imposing any restriction on whether specific combinations of the loci should match the optimum, or that the mean trait itself is at that point. In general, $\bar{z} \neq z_o$, and also $\langle \bar{z} \rangle \neq z_o$.

At the moments, I have found hope to be able to predict the evolution of quantitative characters. The pieces seem to be coming together to answer the fundamental question in quantitative genetics: how is genetic variability maintained? and its related dynamics counterpart: how does genetic variability evolves? I have given only partial answers but the trend is set,

and hopefully sooner than later conclusive answers will come, for which only some details have to be worked out, as explained in the following sub-section.

7.4.1 Overcoming numerical limitations

Unfortunately for the problem dealt with, the numerics are difficult. The main reason for this is that the integrals that we need to perform to compute the observables, involve the product of the per-locus partition functions. These products are often falling in the limits of numerical zero. Of course, since all macroscopics involve the ratio of an integral with the partition function, although each term in itself is small, their ration converges to a finite number. But numerically this is problematic. Most of the time, even such calculations can actually be performed, but the amount of time that the integrator takes to compute them, is enormous. And this grows with the amount of loci. Second, the calculations involve the inversion of the matrix of covariances, which even numerically is very time costly. One stem on the computation can take (depending on the required precision and on the precise values of the parameters) more than an hour. Thus computing a whole trajectory is for practical terms, impossible. Third, if the step for integrating the trajectory are not small enough, then the computations simply diverge.

This can (and will) be solved. There are two methods that can be combined at this point. On the one hand, we can perform a Monte Carlo simulation to perform the integration, using a variant of the Metorpolis-Hastings algorithm (Metropolis et al., 1953; Hastings, 1970). (Notice that a Monte Carlo simulation with the Metropolis-Hastings method *is not* the same as the Monte Carlo method for integrating a function, Press et al., 2007, Ch. 7 p. 397-402) The virtue of this method is that we

do not need to compute the normalization constant of the distribution (i.e. the partition function) so many of the numerical issues are avoided. Notice here two hindrances. First, the integrals are complex, so the methods needs to be adapted for this situation. For this we can expand the macroscopics in their real and complex part and evaluate them separately. But this separation involves further algebra, which at the moments I have not explored.

The second way to solve it, is representing the integrals in terms of a series expansion with respect to $N\sigma$. Then the resulting terms are statistics of directional selection with dominance effects. These statistics are easier to compute, as shown in the previous section. The problem there is that we must include multinomial terms, whose sums are also hard to compute. A possible solution is to sample randomly the multinomial distributions (which is actually faster!) and evaluate the macroscopics at these points.

I have advanced with both method, although there is still some tailoring to be done and implement working versions. For the moments, further insights are disguised in the complexities of the analytic results and in the hindrances to unveil them from the respective numerical computations.

7.5 POSTSCRIPT ON STABILIZING SELECTION

In the mean time between the culmination of this thesis and a day before of sending to print, there has been substantial advance in the statistical mechanics theory regarding stabilizing selection. AS mentioned above, the main limitations are with respect to the computing times. There is a way to approximate the integrands of the partition function in terms of a Gaussian distribution. This approximation is based on the fact that

the partition function of directional selection, is the product of several independent per-locus partition functions (Eq. 7.10). Since these partition functions are in essence characteristic functions, their product is in the limit of large n a Gaussian function, as a consequence of the central limit theorem. The problem is complicated enough, and we are proceeding step by step. So far we have developed the method for arbitrary many loci of equal effects. Following this reasoning, we obtain that the partition function for the polygenic system is

$$\mathbb{Z} = \frac{\left(\tilde{\mathbb{Z}}^0\right)^n}{\sqrt{1+4N\alpha f_0}} \exp\left[\frac{2f_0 N b^2}{1+4N\alpha f_0}\right] . \tag{7.19}$$

From this formula, all the pertinent statistics follow. That is, the observables $\{\langle U\rangle, \langle \bar{z}\rangle, \langle \bar{z}^2\rangle, \langle \nu\rangle\}$, and the other macroscopics of the matrices C and B (Eq. 7.16). All the formulas are expressed in terms of simpler statistics of single locus of a trait without selection, but for which the genetic variance is selected for, namely the statistics generated by the partition function:

$$\mathbb{Z}^0 = \int_0^1 \exp\left[4N\mu\gamma^2 pq\right] (pq)^{4n\mu-1} dp . \tag{7.20}$$

and for which $f_0 = \langle \bar{z}^2\rangle_0 = 2n\gamma(1-\langle\nu\rangle_0)$.

Figure 7.9 shows how well the approximation is even for as few as for three loci, when compared to the exact integrations of the Fourier method.

The most radical test for our approximation is that when selection changes abruptly. For example a sudden shift of the optimum would trigger a quick response of the trait, and a radical reconfiguration of the genetic states. The prediction of the change of the trait mean and of the genetic variance is thus not a trivial task. In turn, our approximation allows to estimate

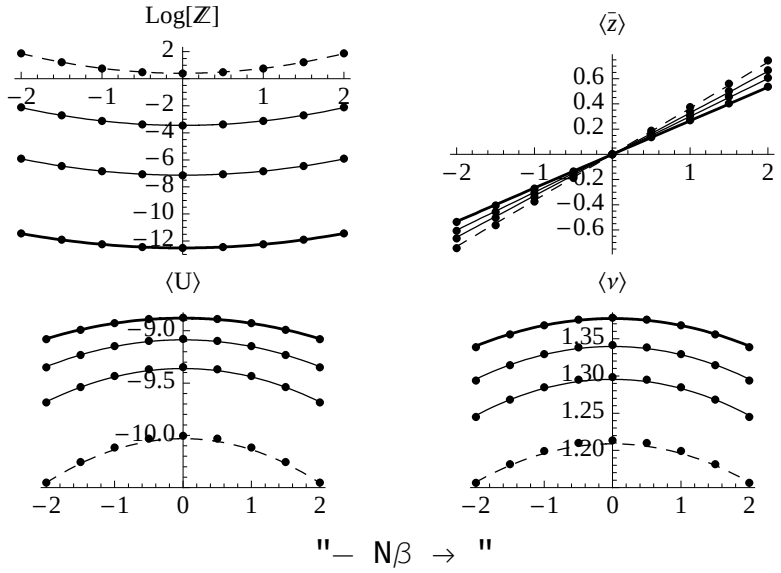

Figure 7.9: Comparing the exact Fourier integration (dots) with the Gaussian approximations for different mutation rates ranging from $N\mu = 0.3$ (dashed lines), 0.5, 0.7 (solid lines), 1.0 (thick solid line). $N\alpha$= -1, $N\sigma = 1$. The trait consists of 3 loci of equal effects = 1

the change of their expectancies, which give robust predictions of their evolutionary course. Figure 7.10 presents a comprehensive analysis of this situation. We compare this response with intensive calculations from the Wright-Fisher model, for distinct numbers of loci. Naturally, the response is quicker for more loci. Unlike the numerical effort required to compute the dynamics, it is reassuring that the precission of the approximation does not seem to depend critically on this number, except for very low number of loci (n between 3 and 5), where the covariances in the matrix C have significant deviations resulting from the Gaussian approximation (results not shown). These

deviations are insignificant for higher number of loci ($n > 10$). Yet the predictions of the macroscopics are in very good agreement with the numerical expectations fromt he Wright-Fisher Model, even for n as low as 4.

However, the change in the genetic variance is very low. In most cases, this change would be so tiny that it would pass inadverted in any practical situation. As it is shown in the previous figure, even when there are conspicuous changes in the mean trait, the changes in the genetic variance are minimal, less than 1% in all cases. (This by the way makes it not only hard but to some extenct pointless to attempt to have an accurate averaging from numerical realizations, and more critically from experiments). This should be compared with the variance from genetic drift fluctuations. Thus it is safer to compare the variance measures of the genetic variance (Var(ν) in an ensemble of populations (realizations) because these are more robust, and is a much more clear cut predition from the SM approach, and thus a way for falsification. After all, there it is meaningless to aim to predict such small changes in genetic variance when we would need an unrealistic number of populations to observe it. For the examples of Fig. 7.10, the fluctuations by drift are so big compared to the range of change of genetic variance, that (a) they completely mask these changes even when averaged for 10^4 realizations, a number of population replicas that is not only unrealistic to achieve even in experiments with micro-organisms, but barely enough to reveal that there is a change, with still an elevated degree of uncertainly, most specially for many loci (Table 7.1).

We can think also of examples where the strenght of selection changes, but without shifting the optimal phenotype. That is, deviations from the optimum become more critical. We can think for example of populations finely adapting to exploiting a particular resource, where a phenotype deviating from the opti-

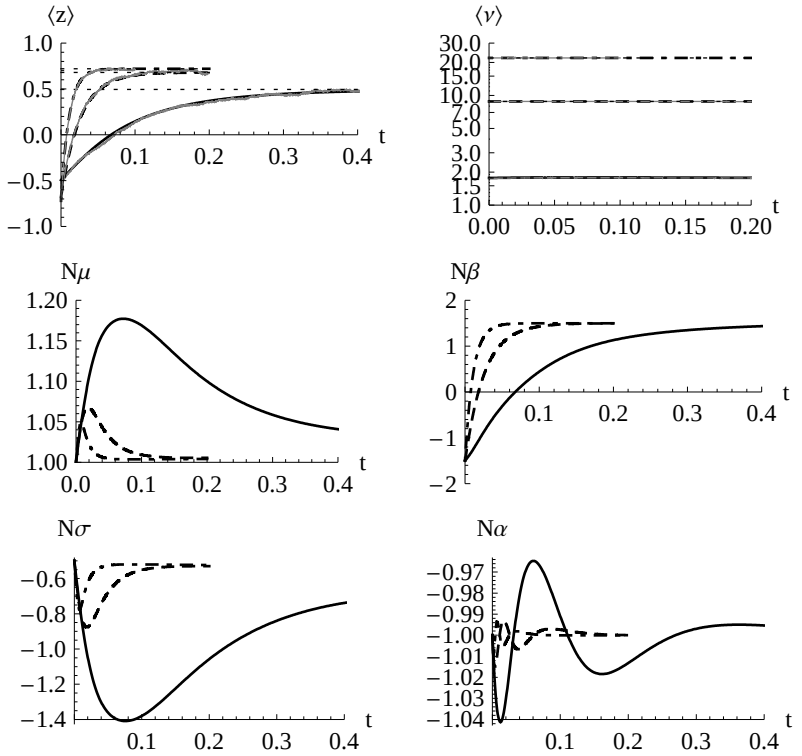

Figure 7.10: Evolution of a polygenic trait under stabilizing selection. Evolutionary response of the trait and genetic variance (top row), and of the local variables (mid and bottom row). Approximations for 4 (solid lines), 20 (dashed lines) and 50 (dot-dashed lines) are shown. The thin lines are averages of the Wright-Fisher model employing 500, 1000, and 5000 replicas respectively. The initial state is given by $N\mu = 1.0, N\beta = -1.5, N\sigma = -1.0, N\alpha = -1.0$, (equivalent to the optimum at $z_o = -1/2$), and $N\beta$ is changed to 1.5 (equivalent to a shift in the optimum to $z_o = -1$). The loci have equal unit effects.

mum (e.g. beak size in the Darwin finches) has less success in exploiting their main resources. The strength of selection would

Table 7.1: Fluctuations in the genetic variance. Traits with distinct numbers of loci (n, first column) show considerable fluctuations in the genetic variance due to drift. Second column, V_ν: statistical mechanical variance of the genetic variance. Third column, $\hat{V}_\nu$: numerical variance form the genetic variance (as in Fig. 7.10). These fluctuations are typically higher than the range of change of the genetic variance (fourth column).

n	V_ν	$\hat{V}_\nu$	Range
2	2.210^{-2}	1.910^{-2} (a)	210^{-2}
20	1.010^{-1}	1.010^{-1} (b)	610^{-3}
50	2.610^{-1}	2.510^{-1} (c)	310^{-3}

a) Variance from 500 replicas.
b) Variance from 2,000 replicas.
c) Variance from 5,000 replicas.

be mediated by a long number of factors, competition, predators, availability of the resource, time allocated to harvest, etc. Any of these factors coul alter the strength of selection without modifying the optimum in a significant way. In such cases, genetic variance changes radically. Some situations that we have successfully tried are when selection becomes disruptive, when selection ceases, or when it is intensified, in all cases without affecting the optimum. The range of change of $\langle\nu\rangle$ is beyond fluctuation by drift and hence the statistical mechanical method predicts accurately the evolutionary trend (data not shown).

As a last point, I would like to briefly comment on the relevance of the previous results. Although it might seem that dealing with the problem of equal effects lacks realism, it provides on the other hand a good tool to understand evolutionary mechanisms. A first example is to understand the dynamics of the microscopic reconfigurations (e.g. following a moving op-

timum; Bürger (2000, pp. 324–331);Jones et al. (2004); Kopp and Hermisson (2007). The allele combinations that are best fit to a given optimum are very sensitive and not linear with the value of such optimum. A small change in that optimal value may involve very different genetic states. The process that allow these changes are the jumps from peaks to peaks in the allele frequencies space (often confused with the fitness landscape, see Ch. 5). The properties of these jumps are in itself a complicated research subject (Barton, 1989; Nagylaki, 1989; Gavrilets and de Jong, 1993; Gavrilets and Hastings, 1993; Coyne et al., 1997; Rogers, 2003; Willensdorfer and Bürger, 2003). Treating the case of equal allelic effects allows to simplify very much the microscopic configurations and disentangle the details of the mechanisms. Because the statistical mechanics allows to have a relatively simple description of the evolutionary dynamics, we can then analyze these situations. A second example that is benefited from assuming equal effects is the contribution of genetic drift to the quantitative variation. Genetic variance evolves erratically due to (a) the rough path in the allele frequencies space, and (b) very unfrequent alleles that sweep in the population. Under unequal effects, these paths are smoothed, thus the contribution due to drift is entangled. Hence assuming equal effects allows to focus on the effects introduced only by drift (Barton et al., 2004; de Brito et al., 2005). A third example is epistatic effects. Again, the complications that epistasis introduce in the response to selection are on the one hand obscure with respect to their contribution to (cryptic) genetic variance (Kondrashov and Turelli, 1992; Gavrilets and de Jong, 1993; Carter et al., 2005; Beerenwinkel et al., 2007; Yukilevich et al., 2008), and on the other hand present a strong non-linear component that is amplified by the presence of alleles evolving due to distinct effective selective strength (Wang et al., 1998; Barton et al., 2004; Roff et al., 2006).

There are of course more examples to be listed, and these are a handful of relevant problems that we are willing to tackle on the framework of statistical mechanics. It is, however not an argument to forget the unequal effects situation. It is not as hard as it might seem on the first look. The Gaussian approximation leading to Eq. 7.19 is just using the central limit theorem. This *does not* require that the variables (allele frequencies) are identically distributed, only independent. Thus it is only a minor complication to extend the analyses for this more realistic situation, which -as that of equal effects- depends on the fact that linkage is not strong. That is the last point which needs to be considered in order to overcome the most stringent limitations between theoretical approaches, and biological reality.